Morphologic and Genomic Analyses of New Isolates Reveal a Second Lineage of Cedratviruses

Rodrigo Araújo Lima Rodrigues,a,b Julien Andreani,a Ana Cláudia dos Santos Pereira Andrade,b Talita Bastos Machado,b Souhila Abdi,a Anthony Levasseur,a Jônatas Santos Abrahão,b Bernard La Scolaa

aAix Marseille Université, IRD, APHM, MEPHI, IHU-Méditerranée Infection, Marseille, France
bLaboratório de Vírus, Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

ABSTRACT Giant viruses have been isolated and characterized in different environments, expanding our knowledge about the biology of these unique microorganisms. In the last 2 years, a new group was discovered, the cedratviruses, currently composed of only two isolates and members of a putative new family, "Pithoviridae," along with previously known pithoviruses. Here we report the isolation and biological and genomic characterization of two novel cedratviruses isolated from samples collected in France and Brazil. Both viruses were isolated using *Acanthamoeba castellanii* as a host cell and exhibit ovoid particles with corks at either extremity of the particle. Curiously, the Brazilian cedratvirus is ~20% smaller and presents a shorter genome of 460,038 bp, coding for fewer proteins than other cedratviruses. In addition, it has a completely asyntenic genome and presents a lower amino acid identity of orthologous genes (~73%). Pangenome analysis comprising the four cedratviruses revealed an increase in the pangenome concomitant with a decrease in the core genome with the addition of the two novel viruses. Finally, phylogenetic analyses clustered the Brazilian virus in a separate branch within the group of cedratviruses, while the French isolate is closer to the previously reported *Cedratvirus lausannensis*. Taking all together, we propose the existence of a second lineage of this emerging viral genus and provide new insights into the biodiversity and ubiquity of these giant viruses.

IMPORTANCE Various giant viruses have been described in recent years, revealing a unique part of the virosphere. A new group among the giant viruses has recently been described, the cedratviruses, which is currently composed of only two isolates. In this paper, we describe two novel cedratviruses isolated from French and Brazilian samples. Biological and genomic analyses showed viruses with different particle sizes, genome lengths, and architecture, revealing the existence of a second lineage of this new group of giant viruses. Our results provide new insights into the biodiversity of cedratviruses and highlight the importance of ongoing efforts to prospect for and characterize new giant viruses.

KEYWORDS *Cedratvirus*, giant virus, NCLDV, new lineage, virion volume, genome length, pangenome

Viruses are the most abundant microorganisms in the biosphere and present the greatest genetic and morphological diversity among the known biological organisms (1, 2). Different groups of viruses have specific characteristics that define them as a group, such as capsid structure (e.g., icosahedral and helical) and type of genome (e.g., double-stranded DNA [dsDNA] and single-stranded RNA negative sense [ssRNA−]), which implicate differences in the life cycles and evolution of these viruses. Moreover, remarkable differences can be seen in the virion volumes and genome

lengths of viruses, exhibiting a difference of 4 orders of magnitude in the former and ranging from 1.2 kbp to 2,500 kbp in the latter (3). Despite these differences, the virion sizes and genome lengths of viruses display an allometric relationship independent of the type of capsid or genetic material, the ebolaviruses (ssRNA−) being the only exception to this scaling law described to date (3). This relationship has also been observed for some giant viruses such as mimivirus and pandoravirus (both dsDNA).

The giant viruses were discovered at the beginning of the last decade with the description of mimiviruses, revealing a new world within the virosphere (4, 5). These viruses replicate in free-living amoebas of the *Acanthamoeba* genus, although other protists have been associated with giant viruses, such as *Cafeteria roenbergensis* (6) and *Bodo saltans* (7). The discovery of viruses with gigantic particles (>500 nm) and the presence of genes considered hallmarks of the cellular world (e.g., those encoding aminoacyl-tRNA synthetases and translation factors) broke many paradigms of classical virology and reopened a hot debate about the origin and evolution of viruses (8–13). In subsequent years, other giant viruses have been isolated and characterized (14, 15), thus expanding the group of nucleocytoplasmic large DNA viruses (NCLDVs)—the proposed "Megavirales" order (16).

Among these new viruses, the pithoviruses drew attention due to their huge size (~1.5 μm) and relatively "small" genomes (~610 kbp) (17, 18), which suggest that those viruses do not fit the scaling law observed for other viruses, although no analysis has been performed in this regard to date. The same case would appear to apply to the cedratviruses, a new group of recently described viruses, with only two members characterized thus far, *Cedratvirus A11* (19), and *Cedratvirus lausannensis* (20). These viruses have an ovoid particle of about 1.0 μm and possess a circular dsDNA genome of ~590 kbp. These viruses are related to the pithoviruses but have two corks in the viral particle, one at either extremity, rather than the single one displayed by pithoviruses. Recently, it has been reported that some pithovirus particles can have complex alternative shapes and sometimes have two corks, as observed for cedratviruses (21). These viruses replicate in *Acanthamoeba* sp., entering the cells by phagocytosis. The genome is released through the cork, and an eclipse phase is established, followed by the formation of an electron-lucent viral factory, where a complex morphogenesis occurs (19, 20, 22). It is possible that there is a nuclear phase during the replication of cedratviruses, since the host nucleus remains intact during the viral cycle, although further investigation into this aspect needs to be performed (22). After ~12 h of infection, mature viral particles are released by cell lysis. The real extent of the diversity of cedratviruses is currently unknown, and the discovery of new members of this group could bring valuable information about it.

Here we describe the isolation and biological and genomic analyses of two new cedratviruses, one from samples collected in France and a second from samples collected in Brazil, which have morphological and genomic features distinct from those of the previously known cedratviruses, suggesting the existence of a second lineage among this new group of viruses. The discovery of new cedratviruses in different parts of the world reflects their ubiquity and high diversity and improves our knowledge about these viruses, thus reinforcing the importance of continuing to prospect for and biologically/genomically characterize the giant viruses.

RESULTS

New cedratviruses with different virion sizes and genome lengths. In the search for a better understanding of the diversity of giant viruses in different parts of the world, we isolated two new cedratviruses, named *Cedratvirus Zaza IHUMI* and *Brazilian Cedratvirus IHUMI*. Transmission electron microscopy (TEM) analyses revealed viruses with ovoid particles and with corks at either extremity of the particle (Fig. 1A to C; see Fig. S1 posted at http://www.mediterranee-infection.com/article.php?laref=983&titer= morphological-and-genomic-analyses-of-new-isolates-reveal-a-second-lineage-of -cedratviruses), a singular feature of cedratviruses (19, 20). Unlike pithoviruses, cedratviruses usually have two corks, although some alternative shapes with a single cork can
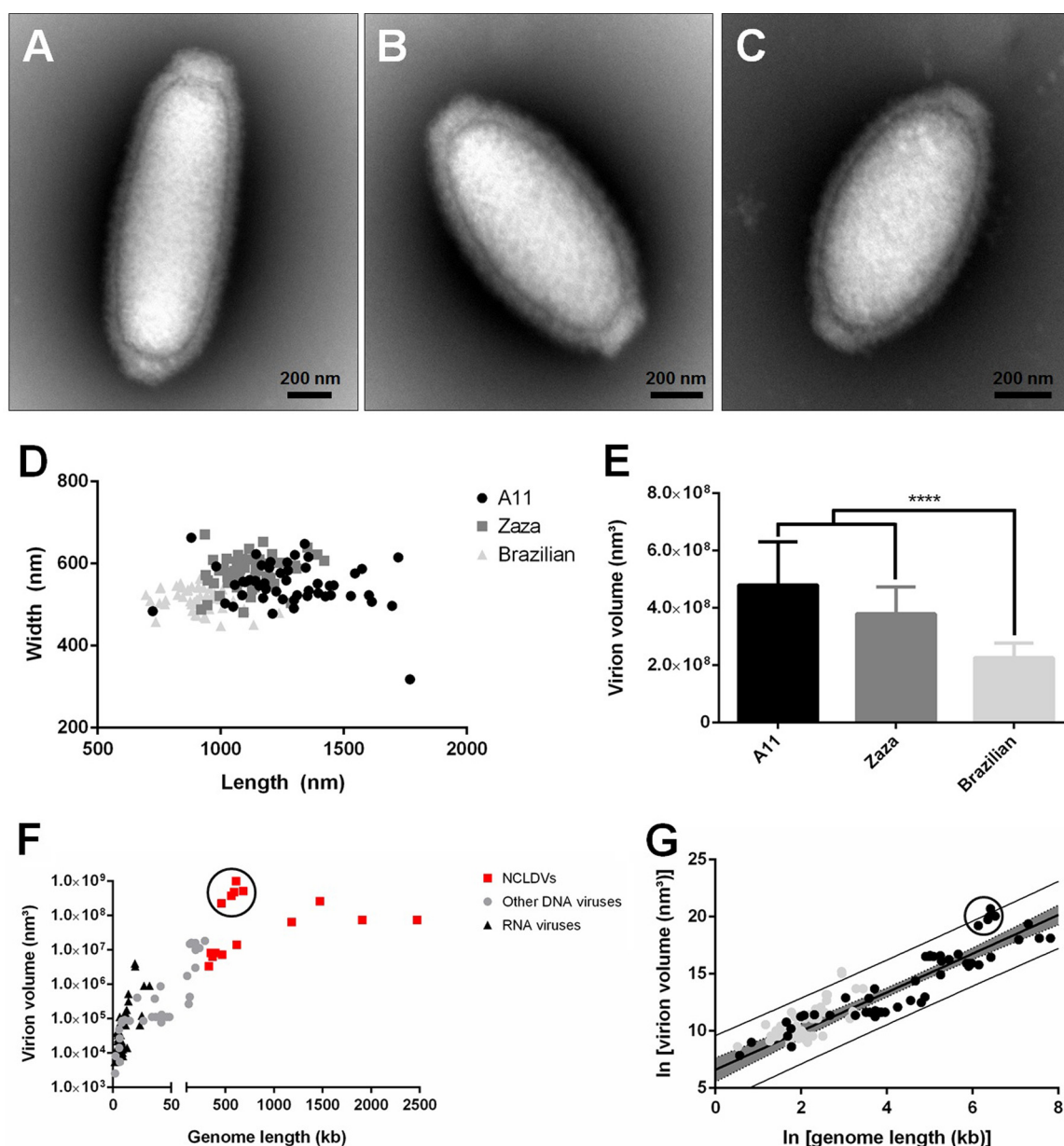
FIG 1 Morphology and volume analysis of new cedratviruses. (A to C) Negative-staining images exhibiting the characteristic ovoid shape and the presence of two corks in the particles of cedratvirus A11, cedratvirus Zaza IHUMI, and Brazilian cedratvirus IHUMI, respectively. Scale bars are indicated on each panel. (D) Length and width of 50 particles of each cedratvirus. Each point represents a single particle analyzed by using ImageJ software. (E) Volumes of different cedratviruses based on the analyses of 50 individual particles, indicating that Brazilian cedratvirus IHUMI has a significantly smaller volume than the other viruses. (F) Relationship between genome length and virion volume for different DNA and RNA viruses. Black circles highlight the pithoviruses and cedratviruses. (G) Relationship between genome length and virion volume for different viruses. The solid black line marks the linear regression between ln-ln-transformed data. The gray area represents the 95% confidence interval for the linear regression line. Black circles highlight the pithoviruses and cedratviruses. The outer gray lines represent the 95% prediction interval, within which we expect 95% of virion volume to lie for a given genome size. ****, $P < 0.0001$ (ANOVA).

be discerned, such as in pithoviruses. Moreover, as recently described for pithoviruses (21), we also observed membranous structures in empty particles in some negative-staining images (see Fig. S2 posted at the URL mentioned above). Cedratvirus Zaza IHUMI particles have a mean size of 1,110 nm (range, 921 to 1,420 nm) in length and 580 nm (range, 481 to 671 nm) in width, values closer to those observed for cedratvirus A11 (1,280 nm by 550 nm), while the Brazilian cedratvirus IHUMI particle is smaller, with the particle displaying a mean size of 910 nm (696 to 1,237 nm) in length and 510 nm

TABLE 1 Main genomic characteristics of known cedratviruses[a]

Virus	Mean particle length × width (nm)	Genome size (bp)	GC content (%)	No. of predicted proteins	Coding density (%)
Cedratvirus A11	1,280 × 550	589,068	42.6	574	78.5
Cedratvirus lausannensis	~1,000 × 500	575,161	42.8	643	83
Cedratvirus Zaza IHUMI	1,110 × 580	560,887	42.7	636	84.3
Brazilian cedratvirus IHUMI	910 × 510	460,038	42.9	533	84.3

[a]All these viruses showed ovoid, double-cork morphology, and none had tRNA.

(448 to 563 nm) in width (Fig. 1D). The difference in the particle size reflects the difference in virion volume, in that the Brazilian Cedratvirus IHUMI has the smallest volume (2.26×10^8 nm³) among the cedratviruses analyzed, significantly smaller than cedratvirus A11 (4.8×10^8 nm³) and cedratvirus Zaza IHUMI (3.79×10^8 nm³) ($P < 0.0001$) (Fig. 1E). Despite these physical differences, the replication cycle of the Brazilian isolate is similar to those previously observed for other cedratviruses, exhibiting the same infection profile (see Fig. S1 posted at the URL mentioned above). The virus enters the host through phagocytosis and releases the capsid content into the cytoplasm, establishing an eclipse phase 2 h after infection. A viral factory is formed in the cytoplasm, where morphogenesis occurs, and mature virions are released 12 h after infection.

In addition to the size of the particles, the new cedratviruses have different genome lengths. Cedratvirus Zaza IHUMI has a genome of 560,887 bp coding for 636 proteins, while the Brazilian cedratvirus IHUMI has a genome of 460,038 bp coding for 533 proteins. Despite the remarkable difference in the length and numbers of predicted open reading frames (ORFs) in the genomes of the new viruses, both exhibit a circular dsDNA genome with the same coding density (84.3%) and very similar G+C contents, 42.7% and 42.9% for cedratvirus Zaza IHUMI and Brazilian cedratvirus IHUMI, respectively (Table 1). It is noteworthy that the Brazilian isolate is the smallest cedratvirus described to date and also has the smallest genome among representatives of this new group of viruses.

Cedratviruses and pithoviruses: exceptions to the allometric scaling law? The fact that the new cedratviruses exhibit different genome lengths led us to analyze the relationship between the genome length and volume size of different NCLDVs, in order to check whether they lie within the prediction interval and are thus in line with the allometric scaling law, as observed for other groups of viruses (3). We calculated the volume of 15 different viruses, including mimiviruses, marseilleviruses, pithoviruses, cedratviruses, faustovirus, kaumoebavirus, pacmanvirus, phycodnavirus, and iridovirus. The volume was calculated by considering the dimensions of viral particles resulting from previous studies using cryo-EM or negative-staining methods (see Table S1 posted at http://www.mediterranee-infection.com/article.php?laref=983&titer=morphological-and-genomic-analyses-of-new-isolates-reveal-a-second-lineage-of-cedratviruses), with the exception of the cedratviruses, for which the volume used in the analysis was the mean volume obtained from the analyses of 50 viral particles using the negative-staining method. Data concerning all other viruses were obtained from previous studies (3).

The volume of the viruses varied by 5 orders of magnitude, with porcine circovirus 1 displaying the smallest volume of the viruses under consideration (2.5×10^3 nm³), and pithovirus sibericum presenting the largest volume (9.9×10^8 nm³). Regarding genome length, this ranged from 1.76 kbp (porcine circovirus 1) to 2,474 kbp (pandoravirus salinus) (Fig. 1F). Considering only the volumes of the NCLDVs calculated in this study, volumes ranged from 3.31×10^6 nm³ (chilo iridescent virus) to 9.9×10^8 nm³ (pithovirus sibericum) (see Table S1 posted at http://www.mediterranee-infection.com/article.php?laref=983&titer=morphological-and-genomic-analyses-of-new-isolates-reveal-a-second-lineage-of-cedratviruses).

Plotting the new data on NCLDVs alongside other viruses on a log-log scale, the linear relationship is maintained ($P < 0.0001$, $R^2 = 0.83$, slope = 1.58), with values even more stringent than those previously reported (3). The vast majority of viruses fall within the 95% prediction interval, indicating that almost all viruses follow the allo-

metric scaling law for volume size and genome length, i.e., the larger the volume size of a viral particle, the longer the genome enclosure by the virus (Fig. 1G). Curiously, cedratviruses and pithoviruses are at the limit of the prediction level, with pithovirus sibericum actually outside the interval. The same was observed when considering only dsDNA viruses (data not shown). This suggests that the putative "Pithoviridae" family could be the first dsDNA group of viruses that does not conform to the allometric scaling law, along with ebolaviruses (ssRNA−). It is notable that although cedratviruses and pithoviruses appear to be exceptions to this scaling law, this appears to be true only when comparing group of viruses, since a virus with a larger volume (e.g., cedratvirus A11) has a longer genome than does a virus displaying a smaller volume, as verified for the Brazilian cedratvirus IHUMI.

Genome comparison of new cedratviruses. The cedratvirus Zaza IHUMI genome exhibit 636 genes, of which 313 (49.2%) code for proteins with no known function (hypothetical proteins). Of these, three had no hits in all searched databases and were considered ORFans (proteins that were longer than 100 amino acids and with no hits in any database). Regarding Brazilian cedratvirus IHUMI, 269 of its 533 predicted genes (50.5%) have no known function and 11 are considered to be ORFans. Among the ORFs with known functions, the presence of genes related to the metabolism of nucleic acids (e.g., those coding for DNA polymerase, DNA-dependent RNA polymerase, helicases, nucleases, DNA repair proteins) and transcription process (e.g., TFIIB initiation factor, TFIIS elongation factor, viral transcription late factor 3) was observed. Moreover, we identified 76 ankyrin repeat-containing-domain proteins in the cedratvirus Zaza IHUMI genome, while only 42 were observed in the Brazilian cedratvirus IHUMI genome. No tRNA or aminoacyl-tRNA synthetases were detected in the genomes of the new viruses. Regarding the nucleocytoplasmic virus orthologous group (NCVOG) core genes, we found some conserved genes also present in some other NCLDVs, e.g., those encoding a divergent major capsid protein (NCVOG0022), D5 helicase-primase (NCVOG0023), DNA topoisomerase II (NCVOG0037), ribonucleotide reductase (NCVOG0276 and NCVOG1353), and an mRNA capping enzyme (NCVOG1117) similar to that observed for other cedratviruses (19, 20).

Although the gene content does not exhibit significant differences at first glance, the genome organization of the Brazilian isolate is completely different from that observed for other cedratviruses, being totally asyntenic (Fig. 2). The synteny analysis revealed the presence of conserved and aligned blocks between cedratvirus Zaza IHUMI, cedratvirus A11, and cedratvirus lausannensis, while the same blocks are organized in a different orientation in the Brazilian cedratvirus IHUMI genome. Compared to the genome of the other viruses, the genome of the Brazilian isolate exhibits many inversions and rearrangements of blocks throughout its entire length. Such differences in the genomic architecture among similar viruses are observed among different lineages of mimiviruses (23) and marseilleviruses (24), which led us to consider the existence of a second lineage of cedratviruses, with Brazilian cedratvirus IHUMI being its first member.

In addition, the Brazilian cedratvirus IHUMI amino acid sequences showed lower identity than other cedratviruses (Fig. 3). The orthologous genes of the Brazilian isolate have a mean identity of 73.48% compared to cedratvirus A11, 73.6% compared to cedratvirus lausannensis, and 73.56% compared to cedratvirus Zaza IHUMI (Fig. 3A to C). In contrast, when we compared the orthologous genes from other cedratviruses to one another, we observed an amino acid identity higher than 90%, reaching 95.76% between cedratvirus lausannensis and the new isolate, cedratvirus Zaza IHUMI (Fig. 3D to F). Therefore, not only is the genomic architecture between the Brazilian isolate and the other viruses different, but also amino acid identity is considerably different, reinforcing the existence of a new lineage among the group of cedratviruses.

Pangenome and phylogenetic analyses of cedratviruses. The pangenome analysis of the cedratviruses isolated thus far revealed an increase in the pangenome content with the addition of a gene repertoire by way of the new viruses described in
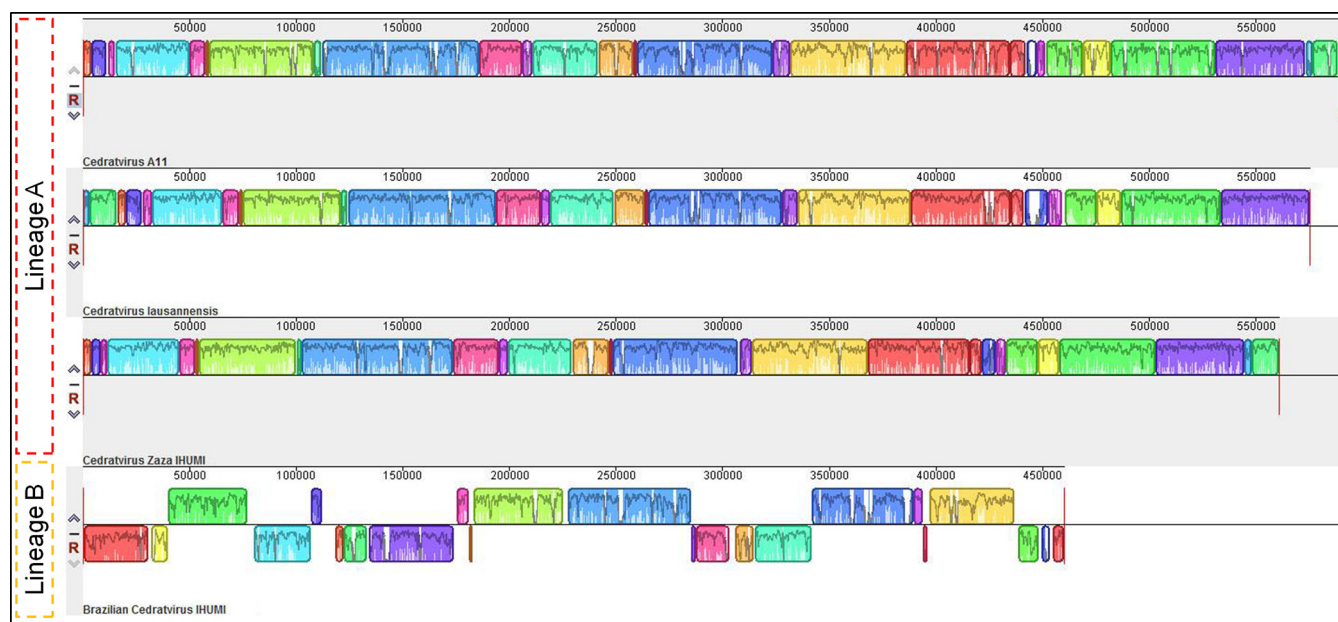
FIG 2 Genome synteny analysis. Schematic genome alignment diagram obtained using the Mauve software package. The analysis was performed using the genome of cedratvirus A11 (NC_032108.1) and cedratvirus lausannensis (LT907979.1), besides the genome sequences of the new isolates. The blocks illustrated above the x axis are in the positive strand (forward sense), while blocks below the x axis are in the negative strand (reverse sense).

this study. A total of 2,386 proteins were grouped into 821 clusters of orthologous genes (COGs) (Fig. 4), including 613 COGs comprising genes for at least two proteins of different cedratviruses. The size of the pangenome content displayed a continuous increase with the addition of the two new viruses, including an increase of 61 new COGs with the addition of the Brazilian cedratvirus IHUMI, even though this virus presented a genome coding for fewer proteins than the other viruses. Furthermore, it is the virus that presents the greatest number of unique COGs (numbering 72), while the others present only 59 (cedratvirus lausannensis), 47 (cedratvirus Zaza IHUMI), and 30 (cedratvirus A11) unique COGs (see Fig. S3 posted at http://www.mediterranee -infection.com/article.php?laref=983&titer=morphological-and-genomic-analyses-of -new-isolates-reveal-a-second-lineage-of-cedratviruses). However, the most remarkable point is the existence of a break in the curve of the core genome content when genes of the Brazilian isolate are added (−102), reaching a total of 386 COGs for this proposed viral genus (Fig. 4). These data suggest that different lineages of cedratviruses could contribute to a slight increase in the pangenome and could share a reduced core gene set.

To better understand the evolutionary relationship between the new cedratviruses and other members of the proposed Megavirales order, we performed phylogenetic analyses based on different NCLDV genes (NCVOGs) including those coding for the family B DNA polymerase (NCVOG0038) (Fig. 5), the major capsid protein (NCVOG0022), the DNA-dependent RNA polymerase subunit 1 (NCVOG0274), and the VV-A18 helicase (NCVOG0508) (see Fig. S4 posted at http://www.mediterranee-infection.com/article.php ?laref=983&titer=morphological-and-genomic-analyses-of-new-isolates-reveal-a -second-lineage-of-cedratviruses). Moreover, we performed additional phylogenetic analyses using the D6/D11 helicase (NCVOG0031), DNA repair exonuclease (NCVOG0308), Flap endonuclease (NCVOG1060), and ATP-dependent DNA ligase (NCVOG0034), focusing on the cedratviruses and closer viral groups, i.e., marseille-viruses and irido/ascoviruses (see Fig. S5 posted at the URL mentioned above). Phylogenetic trees recurrently clustered the new isolates alongside previously described cedratviruses, pithoviruses, and orpheovirus. Furthermore, all trees based on the core genes showed the cedratvirus Zaza IHUMI as being closer to cedratvirus lausannensis and cedratvirus A11 and the Brazilian isolate being in a branch distant
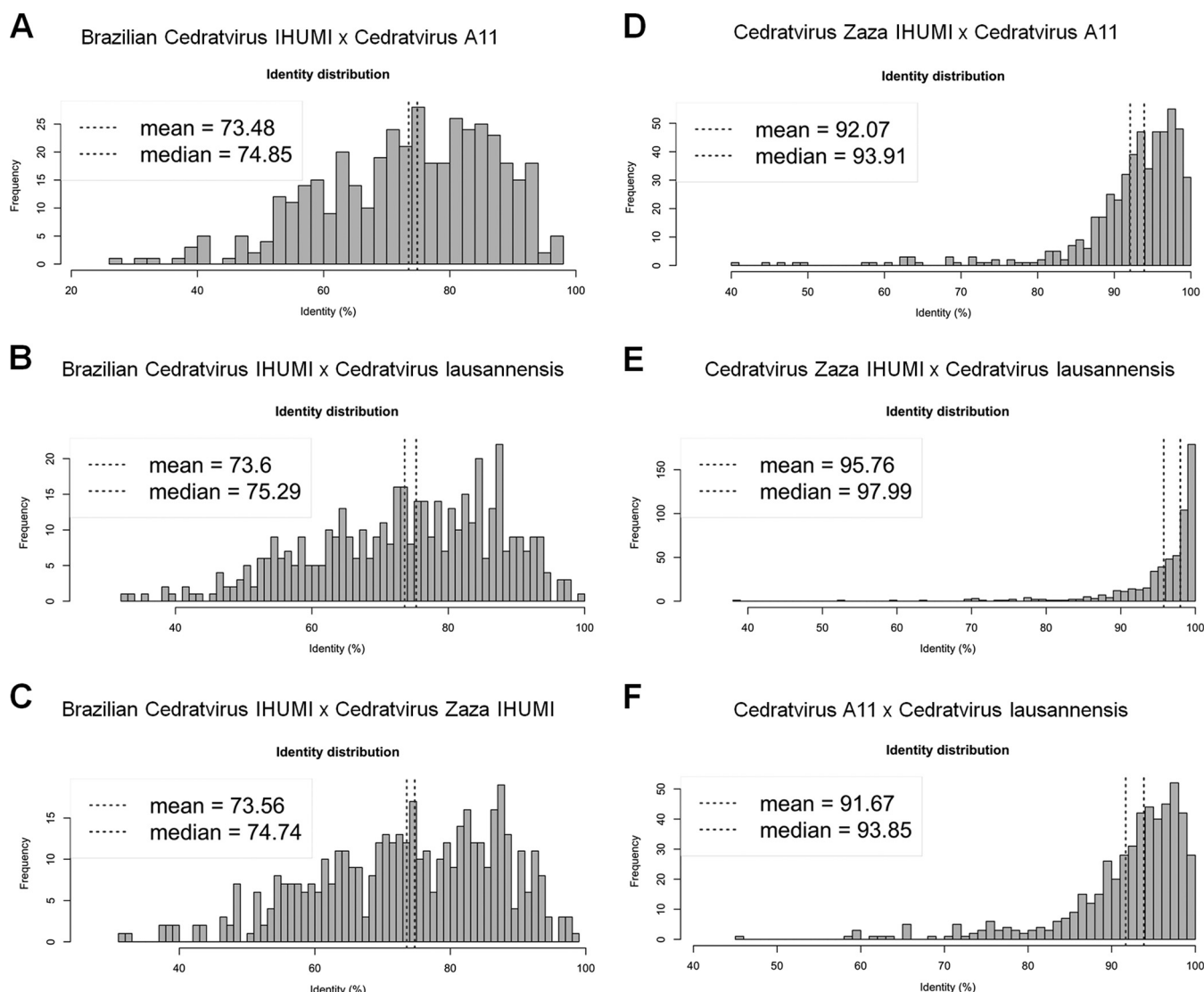
FIG 3 Average amino acid identity. In this analysis, estimates were reached using both best hits (one-way AAI) and reciprocal best hits (two-way AAI) between two data sets of proteins from different cedratviruses. Plots A to C demonstrate the amino acid comparisons between Brazilian cedratvirus IHUMI and other cedratviruses; plots D and E compare cedratvirus Zaza IHUMI and previously known cedratviruses; plot F compares cedratvirus A11 and cedratvirus lausannensis.

from the other cedratviruses with a bootstrap value of >90, corroborating the existence of a new lineage among cedratviruses. Finally, the putative "Pithoviridae" family is clustered along with marseilleviruses or irido/ascoviruses depending of the gene used, the tree topology not being always congruent. An in-depth phylogenetic analysis must be performed to better establish the phylogenetic relationship among these groups of giant viruses.

DISCUSSION

The isolation of new giant viruses associated with biological and genomic analyses has significantly contributed to broadening our understanding of the diversity, ecology, and evolution of this complex group of viruses. The discovery of pithoviruses (17, 18) and cedratviruses (19, 20) drew particular attention, since these viruses exhibit very large particles constraining relatively short genomes, forming a putative novel viral family among the group of NCLDVs. In this study, we describe the isolation and the biological and genomic analyses of two new members of this group, providing new insights into the biodiversity and evolution of these viruses.

The analyses performed in this study revealed two new viruses with significant
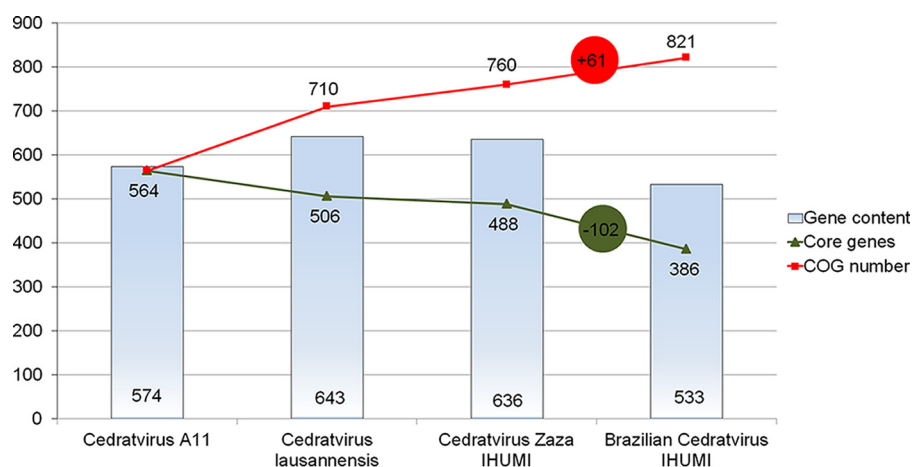
FIG 4 Pangenome (red line) and core genome (green line) sizes of cedratviruses. Numbers at the base of the column refer to the number of genes carried by each virus strain. Numbers at line nodes represent the cumulative COG numbers after the inclusion of a new genome. Numbers in (red and green) circles demonstrate the variation of COGs after the inclusion of the Brazilian cedratvirus IHUMI (proposed new lineage).

structural differences, both physical and genomic. Although the cedratvirus Zaza IHUMI exhibits a particle size and genome length similar to those of other cedratviruses that have been described, analysis of the Brazilian cedratvirus IHUMI revealed a virus with particles that were smaller (~20%) than those of the other viruses of the same group and a considerably smaller genome. By analyzing the relationship between the virion volume and the genome length of viruses, including those from different groups of giant viruses, we noticed that the majority of viruses fall into the allometric scaling law and, curiously, the pithoviruses and cedratviruses are at the limits of the prediction interval. This suggests that these viruses might be exceptions to this scaling law. Since we considered only data from comparable imaging methods (i.e., cryo-EM and negative staining [25, 26]) to calculate the volumes of giant viruses, only a few viruses were analyzed. It is possible that with new, forthcoming structural data on viruses, particularly on giant viruses, it may be discovered that the pithoviruses and cedratviruses definitely fall outside the prediction interval. Indeed, when the virion size data for other amoebal giant viruses (e.g., mimiviruses and marseilleviruses) from TEM images were considered in our analysis, the members of the putative "Pithoviridae" did not fit with the allometric scaling law (data not shown). It is notable that, along with ebolaviruses (*Filoviridae* family), the members of the putative "Pithoviridae" family are the only known viruses that display a massive particle but a "small" genome. Such features raise important questions about what is inside these viral particles. A recent study comparing the internal density of pithoviruses' and mimiviruses' particles demonstrated that the former viruses have three-quarters of the internal density of the latter, suggesting that the pithoviruses may carry macromolecules other than nucleic acids inside the particles (21). The same would appear to be the case for the cedratviruses, but further studies are needed to define exactly which macromolecules could be carried by those viruses.

The fact that the Brazilian isolate has a smaller genome is equally curious. Similar to other cedratviruses, this new virus exhibits only a few repeat zones throughout the genome (data not shown), and we identified the presence of genes also present in other cedratviruses, such as those coding for polymerases, helicases, nucleases, etc. Among the genes with known function, we noticed differences mainly in the quantity of those coding for proteins containing repeat domains, especially coding for ankyrin repeat motifs, as Brazilian cedratvirus IHMU (a virus with a smaller genome) has fewer genes of these category than do other cedratviruses. This is in
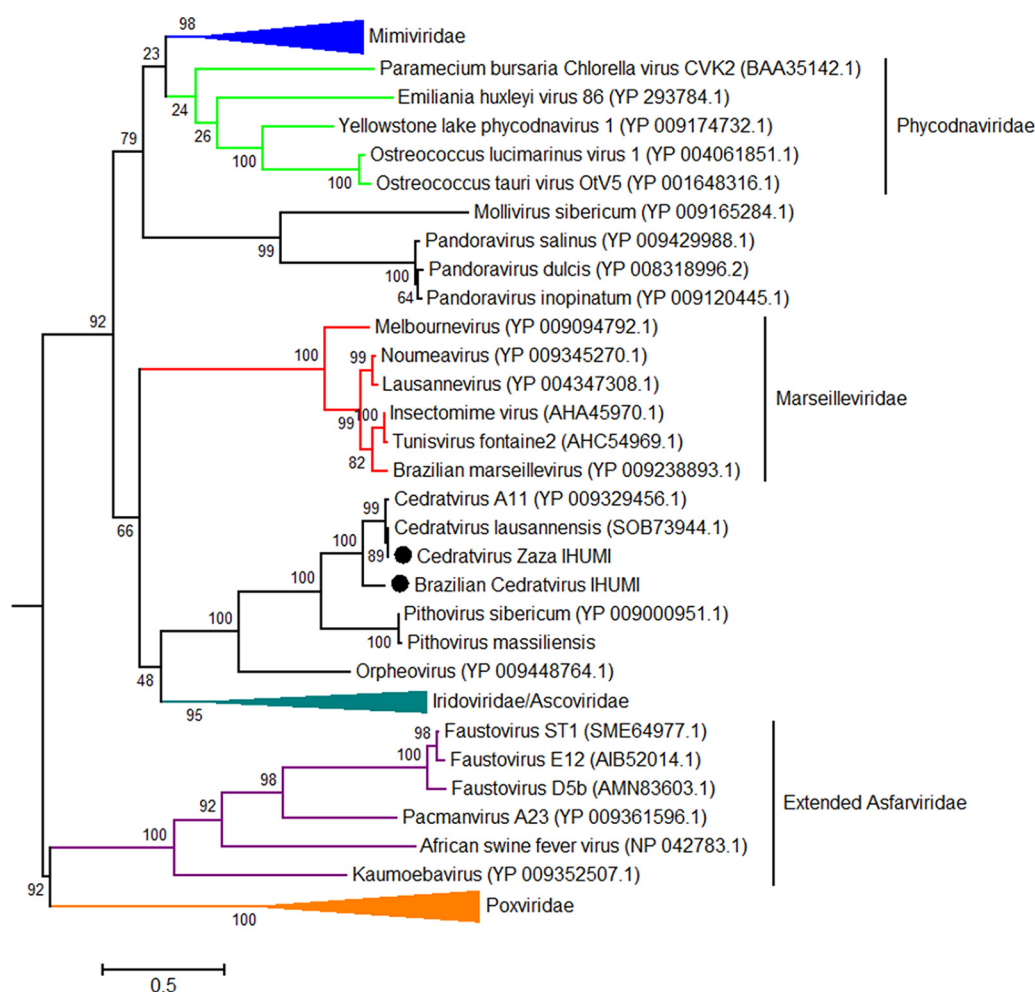
FIG 5 Phylogenetic tree based on DNA polymerase B amino acid sequences of nucleocytoplasmic large DNA viruses (NCLDVs). The tree was constructed using MEGA version 6.0, applying the maximum likelihood method and the JTT model of evolution with 1,000 bootstrap replicates. Colors indicate viral families: blue was used for *Mimiviridae*; green for *Phycodnaviridae*, red for *Marseilleviridae*, navy blue for *Iridoviridae/Ascoviridae*, purple for extended *Asfarviridae*, and orange for *Poxviridae*. The new cedratviruses are highlighted with black circles. The scale bar indicates the rate of evolution.

accordance with the recent analysis conducted by Shukla and colleagues, wherein they demonstrated that in different groups of giant viruses infecting amoebae, the quantity of this class of genes is proportional to the length of the genome (27), which has also been observed for some intracellular bacteria (28). These analyses also seem to apply to viruses within the same group, such as the cedratviruses described here. Taking this into account, it is possible that the Brazilian cedratvirus IHUMI underwent different selective pressures, thus contributing to a different evolutionary history. This would be in accordance with our proposed new lineage within the cedratviruses. Such a proposal is supported by the observation of a completely different genomic architecture between the Brazilian cedratvirus IHUMI and the other viruses, in addition to a significant difference in the amino acid identity of orthologous genes, similar to that observed for members of the *Mimiviridae* and *Marseilleviridae* families (23, 24). In addition, this virus has more exclusive COGs and contributes to an increase in the pangenome with 61 new COGs and, even more strikingly, with the reduction of the core genome by 102 COGs. Finally, phylogenetic analyses based on different core genes of giant viruses clearly clustered the Brazilian isolate in a separate branch from other cedratviruses, therefore reinforcing the existence of a lineage "B" among this new group of viruses.

It is still premature to dive deep into the evolutionary history of cedratviruses, but it is possible that they underwent an accordion-like model of evolution as observed for other giant viruses (29), although new analysis must be performed to confirm this hypothesis. In any case, it is clear that this new, expanding group of viruses deserves attention, and new structural and evolutionary analyses could help to solve some unanswered questions around them.

MATERIALS AND METHODS

Virus isolation, production, and purification. Two novel cedratviruses were isolated by the coculture method as previously described (30), one from an *Alpova* sp. (*Basidiomycota*, *Paxillaceae* family) homogenate in sterile distilled water collected near Toulon, France, and the other from a water sample supplemented with bio-floc, collected in Belo Horizonte, Brazil. The former virus was isolated at the Institut-Hospitalo-Universitaire (IHU) Méditerrannée Infection at Marseille, France, and was named *Cedratvirus Zaza IHUMI*, while the second was isolated in the Laboratürio de Vírus of UFMG at Belo Horizonte, Brazil. The Brazilian isolate was then sent to IHU for further production, genome sequencing, and analysis and was given the name of *Brazilian Cedratvirus IHUMI*. For multiplication of the viruses, *Acanthamoeba castellanii* (strain Neff [ATCC 30010]) was cultured in a 150-cm² cell culture flask with 50 ml of a peptone-yeast extract-glucose (PYG) medium at 28°C. When the flasks contained a fresh monolayer of *A. castellanii*, the PYG medium was replaced by starvation medium (TS). The amoebas were then infected with the isolated virus, and the flasks were kept at 30°C for 72 h. The cell lysates were then collected and centrifuged at 400 × *g* for 10 min to remove amoeba debris. The supernatants were then centrifuged at 6,500 × *g* for 15 min at 4°C, and the pellets were suspended in Page's amoeba saline (PAS) solution. This process was repeated twice. The pellets were suspended in 3 ml of phosphate-buffered saline (PBS) solution, added to a sucrose cushion (50%), and centrifuged at 16,000 × *g* for 15 min at 4°C. The final pellets were suspended in PAS solution.

Characterization of the replicative cycle. In order to study possible differences in the replicative cycle of Brazilian cedratvirus IHUMI, ultrathin sections of infected amoebas were evaluated under TEM and a comparative one-step growth curve was performed. For the microscopy analysis, *A. castellanii* cells were infected with Brazilian cedratvirus IHUMI at a multiplicity of infection (MOI) of 0.01 for 24 h in PYG medium (asynchronous cycle). The cells were then collected and fixed with 2.5% glutaraldehyde in a 0.1 M sodium phosphate buffer for 1 h at room temperature. The amoebas were postfixed with 2% osmium tetroxide and embedded in Epon resin. Ultrathin sections were then analyzed under transmission electron microscopy (Spirit Biotwin FEI, 120 kV). For the one-step growth curves, *A. castellanii* cells were infected with different cedratviruses at an MOI of 10 in TS medium in 24-well microplates. After 30 min of adsorption, the inoculum was removed, and fresh medium was added. The cell and supernatant were collected at different time points and further titrated using the endpoint method (31). The experiment was performed in duplicate.

Virus particle morphometry and volume calculation. For particle morphometry, negative staining was performed on the fixed supernatant from coculture. A total of 5 μl was deposited onto the glow-discharged grid for 20 min at room temperature. The dried grid was contrasted with a small drop of 1% ammonium molybdate for 5 s, and the grid was then observed using a Tecnai G20 electron microscope (FEI, Germany) operated at 200 kV. At least 50 particles of each virus were analyzed using ImageJ software (32). For the volume calculation of cedratvirus particles, we employed the formula for spheroid particles as previously described for ovoid viruses (3), $V = 4/3 \times \pi a^2 c$, where V is the volume of the viral particle, a is the equatorial radius of the spheroid, c is the distance from the center to the pole along the symmetry axis, and π is a constant. The data used for the volume calculation of other NCLDVs were obtained from previous publications, considering data only from cryo-electron microscopy or transmission electron microscopy negative-staining data (see Table S1 posted at http://www.mediterranee-infection.com/article.php?laref=983&titer=morphological-and-genomic-analyses-of-new-isolates-reveal-a-second-lineage-of-cedratviruses). For icosahedral viruses, we used the formula for spherical particles, $V = 4/3 \times \pi r^3$, employing a strategy described elsewhere (3), where r is one-half of the diameter of the virus capsid. For other viruses, we kept the volume data previously calculated by Cui and colleagues (3).

Statistical analysis. We used analysis of variance (ANOVA) to compare the virion volumes of different cedratviruses and linear regression between the natural logarithm of genome length and the natural logarithm of virion volume to test whether the allometric relationship previously described for other viruses (3) also applied to giant viruses of amoebas, which had not previously been evaluated. The statistical analyses were performed by using GraphPad Prism 6.0.

DNA extraction and genome sequencing and assembly. The genomes of the new cedratviruses were extracted using the automated EZ1 virus minikit v.2 (Qiagen GmbH, Hilden, Germany) according to the manufacturer's instructions. DNA was quantified using a Qubit assay with the high-sensitivity kit (Life Technologies, Carlsbad, CA, USA) to 30.3 ng/μl (cedratvirus Zaza IHUMI) and 16 ng/μl (Brazilian cedratvirus IHUMI). DNA was sequenced using MiSeq Technology (Illumina Inc., San Diego, CA, USA) with the paired-end application. DNA was barcoded in order to be mixed with other projects for the Nextera XT DNA sample prep kit (Illumina).

To prepare the paired-end library, dilution was performed to yield 1.0 ng of each genome as input. The "tagmentation" step fragmented and tagged the DNA. Limited-cycle PCR amplification (12 cycles) then completed the tag adapters and introduced dual-index barcodes. The library profile was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies Inc., Santa Clara, CA, USA) with a DNA high-

sensitivity LabChip, and the fragment size was estimated to 1.5 kb. After purification on AMPure XP beads (Beckman Coulter Inc., Fullerton, CA, USA), the libraries were then normalized on specific beads according to the Nextera XT protocol (Illumina). Normalized libraries were pooled for sequencing on the MiSeq. Automated cluster generation and paired-end sequencing with dual index reads were performed in a single 39-hour run in $2\times$ 250 bp. A total of 2.8 Gb of information was obtained from a 277,000/mm² cluster density in the first run with a cluster passing quality control filters of 98.2% (5,333,000 passed filtered clusters). Within this run, the index representation for cedratvirus Zaza IHUMI was determined to be 2.18%. The 149,880 paired-end reads were trimmed and filtered according to the read qualities. Additionally, a total of 7.5 Gb of information was obtained from an 802,000/mm² cluster density in the second run with a cluster passing quality control filters of 96.4% (14,444,000 clusters). Within this run, the index representation for Brazilian cedratvirus IHUMI was determined to be 8.75%. The 1,264,356 paired-end reads were filtered according to the read qualities.

In addition, a run was performed through the MinIon Oxford Nanopore for the Brazilian isolate. The Oxford Nanopore approach was performed on 1D genomic DNA sequencing for the MinIon device using the SQK-LSK108 kit. A library was constructed from 1.5 μg genomic DNA without fragmentation and end repair. Adapters were ligated to both ends of the genomic DNA. After purification on AMPure XP beads (Beckman Coulter Inc., Fullerton, CA, USA), the library was quantified using a Qubit assay with the high-sensitivity kit (Life Technologies, Carlsbad, CA, USA). An amount of 136.8 ng, adapted and tethered as a library, was loaded on the flow cell via the SpotON port. A total of 1,110 active pores were detected for sequencing, and the WIMP workflow was chosen for live bioinformatic analysis. After 4 h and 40 min of run time, the EPI2ME software led to 6,299 classified reads of the Brazilian cedratvirus IHUMI of the 98,601 analyzed reads, with an average length of 2.6 kb.

The sequence reads were assembled *de novo* using the CLC Genomics Workbench v7.5 (http://www.clcbio.com/blog/clc-genomics-workbench-7-5/) for the cedratvirus Zaza IHUMI and hybridSPAdes (33) for the Brazilian cedratvirus IHUMI.

Study of genome organization and genome annotation. Open reading frames were predicted by GeneMarkS (34), and the draft genomes were deposited in NCBI. The tRNA genes were searched using the tRNAscan-SE and ARAGORN software (35, 36). Predicted proteins of fewer than 50 amino acids in length were discarded. A BLASTp search against the NCBI nonredundant (nr) protein sequence database was performed on 5 January 2018. Homology was considered significant if the E value was lower than 1×10^{-3}. A BLASTp search was also computed with the same parameters against the clusters of orthologous groups (COGs) of proteins of the nucleocytoplasmic large DNA virus (known as NCVOGs) (37). In addition, we searched for conserved domains and putative functions of predicted proteins using the online InterProScan software, version 66.0 (https://www.ebi.ac.uk/interpro/search/sequence-search). In addition, predicted ORFs ranging from 50 to 99 amino acids were submitted for tridimensional folding analyses using Phyre2 (38). Proteins ranging from 50 to 99 amino acids in length were discarded if they exhibited no hits either against the BLASTp or against the InterProScan searches or if they exhibited abnormal folding as modeled by Phyre2. Those proteins that were longer than 100 amino acids and with no hits in any database were kept and referred to as ORFans. Finally, the genome annotation was manually revised and curated.

Comparative genomic and pangenome analysis. The genome synteny among cedratviruses was checked using the Mauve program (39) with default parameters. The Proteinortho tool (40) was used to identify orthologous gene sequences based on the reciprocal best hit shared by cedratviruses using an amino acid sequence identity of 30% and sequence coverage of 60% as thresholds. The average amino acid identity (AAI) calculator tool (41) was used to compare identity between orthologous genes from cedratvirus isolates. To estimate the size of the pangenome, their predicted proteins were clustered using the Proteinortho tool (40), applying the same criteria as those given above. We also described pangenome and core genes size variation by stepwise inclusion of each new virus annotation in the pairwise comparisons of the gene contents of the available cedratvirus genome sequences.

Phylogenetic analysis. Phylogenetic reconstructions were based on individual alignment of distinct genes, namely, those encoding the DNA polymerase B family, the DNA-dependent RNA polymerase subunit 1, the VV-A18 helicase, the major capsid protein, the D6/D11 helicase, the Flap endonuclease, the ATP-dependent DNA ligase, and the DNA repair exonuclease. Amino acid sequences were aligned using the Muscle software (42). Phylogenetic trees were built using the MEGA6 software (43), the Jones-Taylor-Thornton (JTT) model for amino acid substitution, and the maximum likelihood method with 1,000 bootstrap replicates.

Accession number(s). Sequences for the draft genomes were deposited in NCBI under the accession numbers LT994651 (Brazilian cedratvirus IHUMI) and LT994652 (cedratvirus Zaza IHUMI).

R.A.L.R. and J.A. performed genomic analyses and wrote the paper; R.A.L.R. performed morphological and phylogenetic analyses; J.A., S.A., A.C.S.P.A., and T.B.M. performed virus isolation and initial biological characterization; A.L., J.S.A., and B.L.S. conceived the study and revised the manuscript.

REFERENCES

1. Suttle CA. 2005. Viruses in the sea. Nature 437:356–361. https://doi.org/10.1038/nature04160.
2. Hulo C, De Castro E, Masson P, Bougueleret L, Bairoch A, Xenarios I, Le Mercier P. 2011. ViralZone: a knowledge resource to understand virus diversity. Nucleic Acids Res 39(Database issue):D576–D582. https://doi.org/10.1093/nar/gkq901.
3. Cui J, Schlub TE, Holmes EC. 2014. An allometric relationship between the genome length and virion volume of viruses. J Virol 88:6403–6410. https://doi.org/10.1128/JVI.00362-14.
4. La Scola B, Audic S, Robert C, Jungang L, de Lamballerie X, Drancourt M, Birtles R, Claverie J-M, Raoult D. 2003. A giant virus in amoebae. Science 299:2033. https://doi.org/10.1126/science.1081867.
5. Colson P, La Scola B, Levasseur A, Caetano-Anollés G, Raoult D. 2017. Mimivirus: leading the way in the discovery of giant viruses of amoebae. Nat Rev Microbiol 15:243–254. https://doi.org/10.1038/nrmicro.2016.197.
6. Fischer MG, Allen MJ, Wilson WH, Suttle CA. 2010. Giant virus with a remarkable complement of genes infects marine zooplankton. Proc Natl Acad Sci U S A 107:19508–19513. https://doi.org/10.1073/pnas.1007615107.
7. Deeg CM, Chow C-ET, Suttle CA. 2018. The kinetoplastid-infecting Bodo saltans virus (BsV), a window into the most abundant giant viruses in the sea. Elife 7:e33014. https://doi.org/10.7554/eLife.33014.
8. Raoult D, Forterre P. 2008. Redefining viruses: lessons from Mimivirus. Nat Rev Microbiol 6:315–319. https://doi.org/10.1038/nrmicro1858.
9. Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, La Scola B, Suzan M, Claverie J-M. 2004. The 1.2-megabase genome sequence of Mimivirus. Science 306:1344–1350. https://doi.org/10.1126/science.1101485.
10. Moreira D, López-García P. 2009. Ten reasons to exclude viruses from the tree of life. Nat Rev Microbiol 7:306–311. https://doi.org/10.1038/nrmicro2108.
11. Raoult D. 2009. There is no such thing as a tree of life (and of course viruses are out!). Nat Rev Microbiol 7:615. https://doi.org/10.1038/nrmicro2108-c6.
12. Koonin EV, Krupovic M, Yutin N. 2015. Evolution of double-stranded DNA viruses of eukaryotes: from bacteriophages to transposons to giant viruses. Ann N Y Acad Sci 1341:10–24. https://doi.org/10.1111/nyas.12728.
13. Nasir A, Caetano-Anollés G. 2015. A phylogenomic data-driven exploration of viral origins and evolution. Sci Adv 1:e1500527. https://doi.org/10.1126/sciadv.1500527.
14. Abergel C, Legendre M, Claverie JM. 2015. The rapidly expanding universe of giant viruses: Mimivirus, Pandoravirus, Pithovirus and Mollivirus. FEMS Microbiol Rev 39:779–796. https://doi.org/10.1093/femsre/fuv037.
15. Colson P, La Scola B, Raoult D. 2017. Giant viruses of amoebae: a journey through innovative research and paradigm changes. Annu Rev Virol 4:61–85. https://doi.org/10.1146/annurev-virology-101416-041816.
16. Colson P, De Lamballerie X, Fournous G, Raoult D. 2012. Reclassification of giant viruses composing a fourth domain of life in the new order Megavirales. Intervirology 55:321–332. https://doi.org/10.1159/000336562.
17. Legendre M, Bartoli J, Shmakova L, Jeudy S, Labadie K, Adrait A, Lescot M, Poirot O, Bertaux L, Bruley C, Coute Y, Rivkina E, Abergel C, Claverie J-M. 2014. Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. Proc Natl Acad Sci U S A 111:4274–4279. https://doi.org/10.1073/pnas.1320670111.
18. Levasseur A, Andreani J, Delerce J, Bou Khalil J, Robert C, La Scola B, Raoult D. 2016. Comparison of a modern and fossil Pithovirus reveals its genetic conservation and evolution. Genome Biol Evol 8:2333–2339. https://doi.org/10.1093/gbe/evw153.
19. Andreani J, Aherfi S, Khalil JYB, Di Pinto F, Bitam I, Raoult D, Colson P, La Scola B. 2016. Cedratvirus, a double-cork structured giant virus, is a distant relative of pithoviruses. Viruses 8:1–11. https://doi.org/10.3390/v8110300.
20. Bertelli C, Mueller L, Thomas V, Pillonel T, Jacquier N, Greub G. 2017. Cedratvirus lausannensis—digging into Pithoviridae diversity. Environ Microbiol 19:4022–4034. https://doi.org/10.1111/1462-2920.13813.
21. Okamoto K, Miyazaki N, Song C, Maia FRNC, Reddy HKN, Abergel C, Claverie JM, Hajdu J, Svenda M, Murata K. 2017. Structural variability and complexity of the giant Pithovirus sibericum particle revealed by high-voltage electron cryo-tomography and energy-filtered electron cryo-microscopy. Sci Rep 7:13291. https://doi.org/10.1038/s41598-017-13390-4.
22. Silva LKDS, Andrade ACDSP, Dornas FP, Rodrigues RAL, Arantes T, Kroon EG, Bonjardim CA, Abrahaõ JS. 2018. Cedratvirus getuliensis replication cycle: an in-depth morphological analysis. Sci Rep 8:4000. https://doi.org/10.1038/s41598-018-22398-3.
23. Assis FL, Franco-Luiz APM, dos Santos RN, Campos FS, Dornas FP, Borato PVM, Franco AC, Abrahao JS, Colson P, La Scola B. 2017. Genome characterization of the first mimiviruses of lineage C isolated in Brazil. Front Microbiol 8:2562. https://doi.org/10.3389/fmicb.2017.02562.
24. Dornas FP, Assis FL, Aherfi S, Arantes T, Abrahão JS, Colson P, La Scola B. 2016. A Brazilian marseillevirus is the founding member of a lineage in family marseilleviridae. Viruses 8(3):76. https://doi.org/10.3390/v8030076.
25. Hoenger A, Aebi U. 1996. 3-D reconstructions from ice-embedded and negatively stained biomacromolecular assemblies: a critical comparison. J Struct Biol 117:99–116. https://doi.org/10.1006/jsbi.1996.0075.
26. De Carlo S, Harris JR. 2011. Negative staining and cryo-negative staining of macromolecules and viruses for TEM. Micron 42:117–31. https://doi.org/10.1016/j.micron.2010.06.003.
27. Shukla A, Chatterjee A, Kondabagil K. 2018. The number of genes encoding repeat domain-containing proteins positively correlates with genome size in amoebal giant viruses. Virus Evol 4:1–11. https://doi.org/10.1093/ve/vex039.
28. Moliner C, Fournier PE, Raoult D. 2010. Genome analysis of microorganisms living in amoebae reveals a melting pot of evolution. FEMS Microbiol Rev 34:281–94. https://doi.org/10.1111/j.1574-6976.2009.00209.x.
29. Filée J. 2015. Genomic comparison of closely related giant viruses supports an accordion-like model of evolution. Front Microbiol 6:593. https://doi.org/10.3389/fmicb.2015.00593.
30. Khalil JYB, Robert S, Reteno DG, Andreani J, Raoult D, La Scola B. 2016. High-throughput isolation of giant viruses in liquid medium using automated flow cytometry and fluorescence staining. Front Microbiol 7:26. https://doi.org/10.3389/fmicb.2016.00026.
31. Reed LJ, Muench H. 1938. A simple method of estimating fifty per cent endpoints. Am J Hyg 27:493–497.
32. Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9:671–675. https://doi.org/10.1038/nmeth.2089.
33. Antipov D, Korobeynikov A, McLean JS, Pevzner PA. 2016. HybridSPAdes: an algorithm for hybrid assembly of short and long reads. Bioinformatics 32:1009–1015. https://doi.org/10.1093/bioinformatics/btv688.
34. Besemer J. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res 29:2607–2618.
35. Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. Nucleic Acids Res 33(Web Server issue):W686–W689.
36. Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res 32:11–16. https://doi.org/10.1093/nar/gkh152.
37. Yutin N, Wolf YI, Raoult D, Koonin EV. 2009. Eukaryotic large nucleo-cytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. Virol J 6:223. https://doi.org/10.1186/1743-422X-6-223.
38. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10:845–858. https://doi.org/10.1038/nprot.2015.053.
39. Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14:1394–1403. https://doi.org/10.1101/gr.2289704.

40. Lechner M, Findeiß S, Steiner L, Marz M, Stadler PF, Prohaska SJ. 2011. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics 12:124. https://doi.org/10.1186/1471-2105-12-124.

41. Rodriguez-R LM, Konstantinidis KT. 2014. Bypassing cultivation to identify bacterial species. Microbe Mag 9:111–118. https://doi.org/10.1128/microbe.9.111.1.

42. Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113. https://doi.org/10.1186/1471-2105-5-113.

43. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729. https://doi.org/10.1093/molbev/mst197.